

# Chapter 2

## The Dynamics of Productivity Change: A Review of the Bottom-Up Approach

Bert M. Balk

**Abstract** This paper considers the relation between (total factor) productivity measures for lower level production units and aggregates thereof such as industries, sectors, or entire economies. In particular, this paper contains a review of the so-called bottom-up approach, which takes an ensemble of individual production units, be it industries or enterprises, as the fundamental frame of reference. At the level of industries the various forms of shift-share analysis are reviewed. At the level of enterprises the additional features that must be taken into account are entry (birth) and exit (death) of production units.

**Keywords** Producer • Productivity • Aggregation • Decomposition • Shift-share analysis • Bottom-up approach • Index number theory

**JEL code:** C43, O47

### 2.1 Introduction

In a previous article (Balk 2010) I considered the measurement of productivity change for a single, consolidated production unit.<sup>1</sup> The present paper continues by studying an ensemble of such units. The classical form is a so-called sectoral shift-share analysis. The starting point of such an analysis is an ensemble of industries,

---

An extended version of this paper is available at SSRN: <http://ssrn.com/abstract=2585452>.

<sup>1</sup>“Consolidated” means that intra-unit deliveries are netted out. In some parts of the literature this is called “sectoral”. At the economy level, “sectoral” output reduces to GDP plus imports, and “sectoral” intermediate input to imports.

B.M. Balk (✉)

Rotterdam School of Management, Erasmus University, Rotterdam, The Netherlands  
e-mail: [bbalk@rsm.nl](mailto:bbalk@rsm.nl)

according to some industrial classification (such as ISIC or NAICE), at some level of detail. An industry is a set of enterprises<sup>2</sup> engaged in the same or similar kind of activities. In the case of productivity analysis the ensemble is usually confined to industries for which independent measurement of input and output is available. Such an ensemble goes by different names: business sector, market sector, commercial sector, or simply measurable sector. Data are published and/or provided by official statistical agencies.

Let us, by way of example, consider labour productivity, in particular value-added based labour productivity. The output of industry  $k$  at period  $t$  is then measured as real value added  $RVA^{kt}$ ; that is, nominal value added  $VA^{kt}$  (= revenue minus intermediate inputs cost) deflated by a suitable, ideally industry-specific, price index. Real value added is treated as ‘quantity’ of a single commodity, that may or may not be added across the production units belonging to the ensemble studied, and over time. At the input side there is usually given some simple measure of labour input, such as total number of hours worked  $L^{kt}$ ; rougher measures being persons employed or full time equivalents employed. Then labour productivity of industry  $k$  at period  $t$  is defined as  $RVA^{kt}/L^{kt}$ .

In the ensemble the industries are of course not equally important, thus some weights reflecting relative importance,  $\theta^{kt}$ , adding up to 1, are necessary. In the literature there is some discussion as to the precise nature of these weights. Should the weights reflect (nominal) value-added shares  $VA^{kt}/\sum_k VA^{kt}$ ? or real value-added shares  $RVA^{kt}/\sum_k RVA^{kt}$ ? or labour input shares  $L^{kt}/\sum_k L^{kt}$ ? We return to this discussion later on.

Aggregate labour productivity at period  $t$  is then defined as a weighted mean, either arithmetic  $\sum_k \theta^{kt} RVA^{kt}/L^{kt}$  or geometric  $\prod_k (RVA^{kt}/L^{kt})^{\theta^{kt}}$ , and the focus of interest is the development of such a mean over time.<sup>3</sup> There are clearly two main factors here, shifting importance and shifting productivity, and their interaction. The usual product of a shift-share analysis is a table which provides detailed decomposition results by industry and time periods compared. Special interest can be directed thereby to industries which are ICT-intensive, at the input and/or the output side; industries which are particularly open to external trade; industries which are (heavily) regulated; *etcetera*.

Things become only slightly more complicated when value-added based *total* factor productivity is considered. At the input side one now needs per industry and time period nominal capital and labour cost as well as one or more suitable deflators. The outcome is real primary input,  $X_{KL}^{kt}$ , which can be treated as ‘quantity’ of another single commodity. Total factor productivity of industry  $k$  at period  $t$  is then defined

<sup>2</sup>There is no unequivocal naming here. So, instead of enterprises one also speaks of firms, establishments, plants, or kind-of-activity units. The minimum requirement is that realistic annual profit/loss accounts can be compiled.

<sup>3</sup>Curiously, the literature neglects the harmonic mean  $(\sum_k \theta^{kt} (RVA^{kt}/L^{kt})^{-1})^{-1}$ ; but, as will be shown in the extended version of this paper, there are conditions under which this mean materializes as the natural one.

as  $RVA^{kt}/X_{KL}^{kt}$ . The issue of the precise nature of the weights gets some additional complexity, since we now also could contemplate the use of nominal or real cost shares to measure the importance of the various industries.

More complications arise when one wants to base the analysis on *gross-output* based total factor productivity. For the output side of the industries one then needs nominal revenue as well as suitable, industry-specific deflators. For the input side one needs nominal primary and intermediate inputs cost together with suitable deflators. The question of which weights to use is aggravated by the fact that industries deliver to each other, so that part of one industry's output becomes part of another industry's input. Improper weighting can then easily lead to double-counting of productivity effects.

Since the early 1990s an increasing number of statistical agencies made (longitudinal) microdata of enterprises available for research. Economists could now focus their research at production units at the lowest level of aggregation and dispense with the age-old concept of the 'representative firm' that had guided so much theoretical development. At the firm or enterprise level one usually has access to nominal data about output revenue and input cost detailed to various categories, in addition to data about employment and some aspects of financial behaviour. Lowest level quantity data are usually not available, so that industry-level deflators must be used. Also, at the enterprise level the information available is generally insufficient to construct firm-specific capital stock data. Notwithstanding such practical restrictions, microdata research has spawned and is still spawning lots of interesting results. A landmark contribution, including a survey of older results, is Foster et al. (2001). Good surveys were provided by Bartelsman and Doms (2000) and, more recently, Syverson (2011). Recent examples of research are collected in a special issue on firm dynamics of the journal *Structural Change and Economic Dynamics* 23 (2012), 325–402.

Of course, dynamics at the enterprise level is much more impressive than at the industry level, no matter how fine-grained. Thought-provoking features are the growth, decline, birth, and death of production units. Split-ups as well as mergers and acquisitions occur all over the place. All this is exacerbated by the fact that the annual microdata sets are generally coming from (unbalanced, rotating) samples, which implies that any superficial analysis of given datasets is likely to draw inaccurate conclusions.

This paper contains a review and discussion of the so-called *bottom-up* approach, which takes an ensemble of individual production units as the fundamental frame of reference. The theory developed here can be applied to a variety of situations, such as (1) a large company consisting of a number of subsidiaries, (2) an industry consisting of a number of enterprises, or (3) an economy or, more precisely, the 'measurable' part of an economy consisting of a number of industries.

The *top-down* approach is the subject of three other papers, namely Balk (2014, 2015) and Dumagan and Balk (2015). The connection between the two approaches, bottom-up and top-down, is discussed in the extended version of the present paper.

What may the reader expect from this review? Sections 2.2 and 2.3 describe the scenery: a set of production units with their accounting relations, undergoing temporal change. Section 2.4 defines the various measurement devices, in particular productivity indices, levels, and their links. The second half of this section is devoted to a discussion of the gap between theory and practice; that is, what to do when not all the data wanted are accessible? And what are the consequences of approximations?

Aggregate productivity change can be measured in different ways. First, as the development through time of arithmetic means of production-unit-specific productivity levels. Section 2.5 reviews the various decompositions proposed in the extant literature, and concludes with a provisional evaluation. Next, Sect. 2.6 briefly discusses the alternatives which emerge when arithmetic means are replaced by geometric or harmonic means. Section 2.7 discusses the monotonicity “problem”, revolving around the so-called Fox “paradox”: an increase of all the individual productivities not necessarily leads to an increase of aggregate productivity. It is argued that this is not a paradox at all but an essential feature of aggregation. Section 2.8 delves into the foundations of the much-used Olley-Pakes decomposition and distinguishes between valid and fallacious use.

In the bottom-up approach aggregate productivity is some weighted mean of individual, production-unit-specific productivities. There is clearly a lot of choice here: in the productivity measure, in the weights of the units, and in the type of mean. Section 2.9 formulates the problem; the actual connection, however, between the bottom-up and top-down approaches is discussed in the extended version of this paper. Section 2.10 concludes with a summary of the main lessons.

## 2.2 Accounting Identities

We consider an ensemble (or set)  $\mathcal{K}^t$  of consolidated production units,<sup>4</sup> operating during a certain time period  $t$  in a certain country or region. For each unit the KLEMS-Y *ex post* accounting identity in nominal values (or, in current prices) reads

$$C_{KL}^{kt} + C_{EMS}^{kt} + \Pi^{kt} = R^{kt} \quad (k \in \mathcal{K}^t), \quad (2.1)$$

where  $C_{KL}^{kt}$  denotes the primary input cost,  $C_{EMS}^{kt}$  the intermediate inputs cost,  $R^{kt}$  the revenue, and  $\Pi^{kt}$  the profit (defined as remainder). Intermediate inputs cost (on energy, materials, and business services) and revenue concern generally tradeable commodities. It is presupposed that there is some agreed-on commodity classification, such that  $C_{EMS}^{kt}$  and  $R^{kt}$  can be written as sums of quantities times (unit) prices of these commodities. Of course, for any production unit most of these

---

<sup>4</sup>In terms of variables to be defined below, consolidation means that  $C_{EMS}^{kkt} = R^{kkt} = 0$ .

quantities will be zero. It is also presupposed that output prices are available from a market or else can be imputed. Taxes on production are supposed to be allocated to the  $K$  and  $L$  classes.

The commodities in the capital class  $K$  concern owned tangible and intangible assets, organized according to industry, type, and age class. Each production unit uses certain quantities of those assets, and the configuration of assets used is in general unique for the unit. Thus, again, for any production unit most of the asset cells are empty. Prices are defined as unit user costs and, hence, capital input cost  $C_K^{kt}$  is a sum of prices times quantities.

Finally, the commodities in the labour class  $L$  concern detailed types of labour. Though any production unit employs specific persons with certain capabilities, it is usually their hours of work that count. Corresponding prices are hourly wages. Like the capital assets, the persons employed by a certain production unit are unique for that unit. It is presupposed that, wherever necessary, imputations have been made for self-employed workers. Henceforth, labour input cost  $C_L^{kt}$  is a sum of prices times quantities.

Total primary input cost is the sum of capital and labour input cost,  $C_{KL}^{kt} = C_K^{kt} + C_L^{kt}$ . Profit  $\Pi^{kt}$  is the balancing item and thus may be positive, negative, or zero.

The KL-VA accounting identity then reads

$$C_{KL}^{kt} + \Pi^{kt} = R^{kt} - C_{EMS}^{kt} \equiv VA^{kt} \quad (k \in \mathcal{K}^t), \quad (2.2)$$

where  $VA^{kt}$  denotes value added, defined as revenue minus intermediate inputs cost. In this paper it will always be assumed that  $VA^{kt} > 0$ .

We now consider whether the ensemble of production units  $\mathcal{K}^t$  can be considered as a consolidated production unit. Though aggregation basically is addition, adding-up the KLEMS-Y relations over all the units would imply double-counting because of deliveries between units. To see this, it is useful to split intermediate input cost and revenue into two parts, respectively concerning units belonging to the ensemble  $\mathcal{K}^t$  and units belonging to the rest of the world. Thus,

$$C_{EMS}^{kt} = \sum_{k' \in \mathcal{K}^t, k' \neq k} C_{EMS}^{k'kt} + C_{EMS}^{ekt}, \quad (2.3)$$

where  $C_{EMS}^{k'kt}$  is the cost of the intermediate inputs purchased by unit  $k$  from unit  $k'$ , and  $C_{EMS}^{ekt}$  is the cost of the intermediate inputs purchased by unit  $k$  from the world beyond the ensemble  $\mathcal{K}$ . Similarly,

$$R^{kt} = \sum_{k' \in \mathcal{K}^t, k' \neq k} R^{kk't} + R^{ket}, \quad (2.4)$$

where  $R^{kk't}$  is the revenue obtained by unit  $k$  from delivering to unit  $k'$ , and  $R^{ket}$  is the revenue obtained by unit  $k$  from delivering to units outside of  $\mathcal{K}^t$ . Adding up the KLEMS-Y relations (2.1) then delivers

$$\sum_{k \in \mathcal{K}^t} C_{KL}^{kt} + \sum_{k \in \mathcal{K}^t} \sum_{k' \in \mathcal{K}^t, k' \neq k} C_{EMS}^{k'kt} + \sum_{k \in \mathcal{K}^t} C_{EMS}^{ekt} + \sum_{k \in \mathcal{K}^t} \Pi^{kt} = \sum_{k \in \mathcal{K}^t} \sum_{k' \in \mathcal{K}^t, k' \neq k} R^{kk't} + \sum_{k \in \mathcal{K}^t} R^{ket}. \quad (2.5)$$

If for all the tradeable commodities output prices are identical to input prices (which is ensured by National Accounting conventions), then the two intra- $\mathcal{K}^t$ -trade terms cancel, and the foregoing expression reduces to

$$\sum_{k \in \mathcal{K}^t} C_{KL}^{kt} + \sum_{k \in \mathcal{K}^t} C_{EMS}^{ekt} + \sum_{k \in \mathcal{K}^t} \Pi^{kt} = \sum_{k \in \mathcal{K}^t} R^{ket}. \quad (2.6)$$

Recall that capital assets and hours worked are unique for each production unit, which implies that primary input cost may simply be added over the units, without any fear for double-counting. Thus expression (2.6) is the KLEMS-Y accounting relation for the ensemble  $\mathcal{K}^t$ , considered as a consolidated production unit. The corresponding KL-VA relation is then

$$\sum_{k \in \mathcal{K}^t} C_{KL}^{kt} + \sum_{k \in \mathcal{K}^t} \Pi^{kt} = \sum_{k \in \mathcal{K}^t} R^{ket} - \sum_{k \in \mathcal{K}^t} C_{EMS}^{ekt}, \quad (2.7)$$

which can be written as

$$C_{KL}^{\mathcal{K}^t t} + \Pi^{\mathcal{K}^t t} = R^{\mathcal{K}^t t} - C_{EMS}^{\mathcal{K}^t t} \equiv VA^{\mathcal{K}^t t}. \quad (2.8)$$

where  $C_{KL}^{\mathcal{K}^t t} \equiv \sum_{k \in \mathcal{K}^t} C_{KL}^{kt}$ ,  $\Pi^{\mathcal{K}^t t} \equiv \sum_{k \in \mathcal{K}^t} \Pi^{kt}$ ,  $R^{\mathcal{K}^t t} \equiv \sum_{k \in \mathcal{K}^t} R^{ket}$ , and  $C_{EMS}^{\mathcal{K}^t t} \equiv \sum_{k \in \mathcal{K}^t} C_{EMS}^{ekt}$ . One verifies immediately that

$$VA^{\mathcal{K}^t t} = \sum_{k \in \mathcal{K}^t} VA^{kt}. \quad (2.9)$$

The similarity between expressions (2.2) and (2.8), together with the additive relation between all the elements, is the reason why the KL-VA production model is the natural starting point for studying the relation between individual and aggregate measures of productivity change. We will see however that the bottom-up approach basically neglects this framework.

### 2.3 Continuing, Entering, and Exiting Production Units

As indicated in the previous section the superscript  $t$  denotes a time period, the usual unit of measurement being a year. Though data may be available over a longer time span, any comparison is concerned with only two periods: an earlier period

0 (also called base period), and a later period 1 (also called comparison period). These periods may or may not be adjacent. When the production units are industries, then the ensemble  $\mathcal{K}^0$  will usually be the same as  $\mathcal{K}^1$ . But when the production units studied are enterprises, this will in general not hold, and we must distinguish between continuing, exiting, and entering production units. In particular,

$$\mathcal{K}^0 = \mathcal{C}^{01} \cup \mathcal{X}^0 \quad (2.10)$$

$$\mathcal{K}^1 = \mathcal{C}^{01} \cup \mathcal{N}^1, \quad (2.11)$$

where  $\mathcal{C}^{01}$  denotes the set of continuing units (that is, units active in both periods),  $\mathcal{X}^0$  the set of exiting units (active in the base period only), and  $\mathcal{N}^1$  the set of entering units (active in the comparison period only). The sets  $\mathcal{C}^{01}$  and  $\mathcal{X}^0$  are disjunct, as are  $\mathcal{C}^{01}$  and  $\mathcal{N}^1$ .

It is important to observe that in any application the distinction between continuing, entering, and exiting production units depends on the length of the time periods being compared, and on the time span between these periods.

Of course, when the production units studied form a balanced panel, then the sets  $\mathcal{X}^0$  and  $\mathcal{N}^1$  are empty. The same holds for the case where the production units are industries. These two situations will in the sequel be considered as specific cases.

The theory developed in the remainder of this paper is cast in the language of intertemporal comparisons. By redefining 0 and 1 as countries or regions, and conditioning on a certain time period, the following can also be applied to cross-sectional comparisons. There is one big difference, however. Apart from mergers, acquisitions and the like, enterprises have a certain perseverance and can be observed through time. But a certain enterprise cannot exist at the same time in two countries or regions. Hence, in cross-sectional comparisons the lowest-level production units can only be industries, and ‘entering’ or ‘exiting’ units correspond to industries existing in only one of the two countries or regions which are compared.

## 2.4 Productivity Indices and Levels

As explained in Sect. 2.2, the various components of the accounting identity (2.1) are nominal values, that is, sums of prices times quantities. We are primarily interested in their development through time, as measured by ratios. It is assumed that all the detailed price and quantity data, underlying the values, are accessible. This is, of course, the ideal situation, which in practice is not likely to occur. Nevertheless, for conceptual reasons it is good to use this as our starting point. More mundane situations, deviating to a higher or lesser degree from the ideal, will then be considered later.

### 2.4.1 Indices

Using index number theory, each nominal value ratio can be decomposed as a product of two components, one capturing the price effect and the other capturing the quantity effect. Thus, let there be price and quantity indices such that for any two periods  $t$  and  $t'$  the following relations hold:

$$C_{KL}^{kt}/C_{KL}^{kt'} = P_{KL}^k(t, t')Q_{KL}^k(t, t') \quad (2.12)$$

$$C_{EMS}^{kt}/C_{EMS}^{kt'} = P_{EMS}^k(t, t')Q_{EMS}^k(t, t') \quad (2.13)$$

$$R^{kt}/R^{kt'} = P_R^k(t, t')Q_R^k(t, t'). \quad (2.14)$$

Capital cost and labour cost are components of primary input cost, thus it can also be assumed that there are functions such that

$$C_K^{kt}/C_K^{kt'} = P_K^k(t, t')Q_K^k(t, t') \quad (2.15)$$

$$C_L^{kt}/C_L^{kt'} = P_L^k(t, t')Q_L^k(t, t'). \quad (2.16)$$

We are using here the shorthand notation introduced in the earlier article (Balk 2010). All these price and quantity indices are supposed to be, appropriately dimensioned, functions of the prices and quantities at the two periods that play a role in the value ratios; e.g.  $P_L^k(t, t')$  is a labour price index for production unit  $k$ , based on all the types of labour distinguished, comparing hourly wages at the two periods  $t$  and  $t'$ , conditional on hours worked at these periods. These functions are supposed to satisfy some basic axioms ensuring proper behaviour, and, dependent on the time span between  $t$  and  $t'$ , may be direct or chained indices (see Balk 2008). There may or may not exist functional relations between the overall index  $P_{KL}^k(t, t')$  and the subindices  $P_K^k(t, t')$  and  $P_L^k(t, t')$  (or, equivalently, between the overall index  $Q_{KL}^k(t, t')$  and the subindices  $Q_K^k(t, t')$  and  $Q_L^k(t, t')$ ).

The construction of price and quantity indices for value added was discussed in Balk (2010, Appendix B). Thus there are also functions such that

$$VA^{kt}/VA^{kt'} = P_{VA}^k(t, t')Q_{VA}^k(t, t') \quad (2.17)$$

Formally, the relations (2.12), (2.13), (2.14), (2.16) and (2.17) mean that the Product Test is satisfied. Notice that it is not required that all the functional forms of the price and quantity indices be the same. However, the Product Test in combination with the axioms rules out a number of possibilities.

We recall some definitions. The *value-added based total factor productivity index* for period 1 relative to period 0 was defined by Balk (2010) as



$$ITFPROD_{VA}^k(1, 0) \equiv \frac{Q_{VA}^k(1, 0)}{Q_{KL}^k(1, 0)}. \quad (2.18)$$

This index measures the ‘quantity’ change component of value added relative to the quantity change of all the primary inputs. The two main primary input components are capital and labour; both deserve separate attention.

The *value-added based capital productivity index* for period 1 relative to period 0 is defined as

$$IKPROD_{VA}^k(1, 0) \equiv \frac{Q_{VA}^k(1, 0)}{Q_K^k(1, 0)}. \quad (2.19)$$

This index measures the ‘quantity’ change component of value added relative to the quantity change of capital input.

Similarly, the *value-added based labour productivity index* for period 1 relative to period 0 is defined as

$$ILPROD_{VA}^k(1, 0) \equiv \frac{Q_{VA}^k(1, 0)}{Q_L^k(1, 0)}. \quad (2.20)$$

This index measures the ‘quantity’ change component of value added relative to the quantity change of labour input. Recall that the labour quantity index  $Q_L^k(t, t')$  is here defined as an index acting on the prices and quantities of all the types of labour that are being distinguished.

Suppose now that the units of measurement of the various types of labour are in some sense the same; that is, the quantities of all the labour types are measured in hours, or in full-time equivalent jobs, or in some other common unit. Then it makes sense to define the total labour quantity of production unit  $k$  at period  $t$  as

$$L^{kt} \equiv \sum_{n \in L} x_n^{kt}, \quad (2.21)$$

and to use the ratio  $L^{kt}/L^{kt'}$  as quantity index. Formally, this is a Dutot or simple sum quantity index. The ratio of a genuine labour quantity index, i.e. an index based on types of labour,  $Q_L^k(t, t')$ , and the simple sum labour quantity index  $L^{kt}/L^{kt'}$  is an index of *labour quality (or composition)*.

The *value-added based simple labour productivity index* for production unit  $k$ , for period 1 relative to period 0, is defined as

$$ISLPROD_{VA}^k(1, 0) \equiv \frac{Q_{VA}^k(1, 0)}{L^{k1}/L^{k0}}, \quad (2.22)$$

which can then be interpreted as an index of real value added per unit of labour.

### 2.4.2 Levels

As one sees, some ‘level’-language has crept in. The bottom-up approach freely talks about productivity (change) in terms of levels. But what precisely are levels, and what is the relation between levels and indices? Intuitively, indices are just ratios of levels, so that it seems that the difference is merely in the kind of language one prefers. It appears, however, that a closer look is warranted.

For each production unit  $k \in \mathcal{K}^t$  *real value added* is (ideally) defined as

$$RVA^k(t, b) \equiv VA^{kt} / P_{VA}^k(t, b); \quad (2.23)$$

that is, nominal value added at period  $t$  divided by (or, as one says, deflated by) a production-unit- $k$ -specific value-added based price index for period  $t$  relative to a certain reference period  $b$ , where period  $b$  may or may not precede period 0. Notice that this definition tacitly assumes that production unit  $k$ , existing in period  $t$ , also existed or still exists in period  $b$ ; otherwise, deflation by a production-unit- $k$ -specific index would be impossible. When production unit  $k$  does not exist in period  $b$  then for deflation a non-specific index must be used. On the complications thereby we will come back at a later stage.

The foregoing definition implies that

$$RVA^k(b, b) = VA^{kb} / P_{VA}^k(b, b) = VA^{kb}, \quad (2.24)$$

since any price index, whatever its functional form, returns the outcome 1 for the reference period. Thus, at the reference period  $b$ , real value added equals nominal value added.

For example, one easily checks that when  $P_{VA}^k(t, b)$  is a Paasche-type double deflator, then real value added  $RVA^{kt}$  is period  $t$  value added at prices of period  $b$  (recall Balk 2010, Appendix B). The rather intricate form at the left-hand side of expression (2.23) serves to make clear that unlike  $VA^{kt}$ , which is an observable monetary magnitude,  $RVA^k(t, b)$  is *the outcome of a function*. Though the outcome is also monetary, its magnitude depends on the reference period and the deflator chosen.

The first kind of dependence becomes clear by considering  $RVA^k(t, b')$  for some  $b' \neq b$ . One immediately checks that  $RVA^k(t, b') / RVA^k(t, b) = P_{VA}^k(t, b) / P_{VA}^k(t, b')$ , which is a measure of the ( $k$ -specific value-added based) price difference between periods  $b'$  and  $b$ . Put otherwise, real value added depends critically on the price level of the reference period, which is the period for which nominal and real value added coincide.

As to the other dependence, it of course matters whether  $P_{VA}^k(t, b)$  is a Paasche-type or a Laspeyres-type or a Fisher-type double deflator. Here the difference in general increases with increasing the time span between the periods  $t$  and  $b$ .

Another way of looking at real value added is to realize that, by using expression (2.17),  $RVA^k(t, b) = VA^{kb} Q_{VA}^k(t, b)$ . Put otherwise, real value added is a (normalized) quantity index.

Like real value added, *real primary, or capital-and-labour, input*, relative to reference period  $b$ , is (ideally) defined as deflated primary input cost,

$$X_{KL}^k(t, b) \equiv C_{KL}^{kt} / P_{KL}^k(t, b); \quad (2.25)$$

*real capital input*, relative to reference period  $b$ , is (ideally) defined as deflated capital cost,

$$X_K^k(t, b) \equiv C_K^{kt} / P_K^k(t, b); \quad (2.26)$$

and *real labour input*, relative to reference period  $b$ , is (ideally) defined as deflated labour cost,

$$X_L^k(t, b) \equiv C_L^{kt} / P_L^k(t, b), \quad (2.27)$$

Of course, similar observations as above apply to these two definitions. In particular, it is important to note that at the reference period  $b$  real primary input equals nominal input cost,  $X_{KL}^k(b, b) = C_{KL}^{kb}$ , real capital input equals nominal capital cost,  $X_K^k(b, b) = C_K^{kb}$ , and real labour input equals nominal labour cost,  $X_L^k(b, b) = C_L^{kb}$ .

It is important to observe that, whereas nominal values are additive, real values are generally not; that is,  $X_{KL}^k(t, b) \neq X_K^k(t, b) + X_L^k(t, b)$  for  $t \neq b$ . It is easy to see, by combining expressions (2.25), (2.26) and (2.27), that requiring additivity means that the overall price index  $P_{KL}^k(t, b)$  must be a second-stage Paasche index of the two subindices  $P_K^k(t, b)$  and  $P_L^k(t, b)$ . When we are dealing with chained indices it is impossible to satisfy this requirement. An operationally feasible solution was proposed by Balk and Reich (2008).<sup>5</sup>

Using the foregoing building blocks, the *value-added based total factor productivity level* of production unit  $k$  at period  $t$  is defined as real value added divided by real primary input,

$$TFPROD_{VA}^k(t, b) \equiv \frac{RVA^k(t, b)}{X_{KL}^k(t, b)}. \quad (2.28)$$

Notice that numerator as well as denominator are expressed in the same price level, namely that of period  $b$ . Thus  $TFPROD_{VA}^k(t, b)$  is a dimensionless variable.

The foregoing definition immediately implies that at the reference period  $b$  value-added based total factor productivity equals nominal value added divided by nominal primary input cost,  $TFPROD_{VA}^k(b, b) = VA^{kb} / C_{KL}^{kb}$ . Now recall the KL-VA

<sup>5</sup>Of course, a trivial solution would be to use the same deflator for all the nominal values. Such a strategy was proposed for the National Accounts by Durand (2004).

accounting identity (2.2) and assume that profit  $\Pi^{kt}$  is constrained to equal 0 for all production units at all time periods. Then reference period total factor productivity of all production units equals 1,  $TFPROD_{VA}^k(b, b) = 1$  ( $k \in \mathcal{K}^t$ ).

Likewise, the *value-added based labour productivity level* of unit  $k$  at period  $t$  is defined as real value added divided by real labour input,

$$LPROD_{VA}^k(t, b) \equiv \frac{RVA^k(t, b)}{X_L^k(t, b)}. \quad (2.29)$$

This is also a dimensionless variable. For the reference period  $b$  we obtain

$$LPROD_{VA}^k(b, b) = \frac{VA^{kb}}{C_L^{kb}} = \frac{VA^{kb}}{C_{KL}^{kb}} \frac{C_{KL}^{kb}}{C_L^{kb}}. \quad (2.30)$$

Hence, when profit  $\Pi^{kt} = 0$  for all production units at all time periods then production unit  $k$ 's labour productivity at reference period  $b$ ,  $LPROD_{VA}^k(b, b)$  equals  $C_{KL}^{kb}/C_L^{kb}$ . This is the reciprocal of  $k$ 's labour cost share at period  $b$ .

In case the simple sum quantity index is used for labour, one obtains

$$LPROD_{VA}^k(t, b) = \frac{RVA^k(t, b)}{C_L^{kt}/P_L^k(t, b)} = \frac{RVA^k(t, b)}{C_L^{kb}Q_L^k(t, b)} = \frac{RVA^k(t, b)}{(C_L^{kb}/L^{kb})L^{kt}}, \quad (2.31)$$

where subsequently expressions (2.27) and (2.16) were used. The constant in the denominator,  $C_L^{kb}/L^{kb}$ , is the mean price of a unit of labour at reference period  $b$ .

The *simple value-added based labour productivity level* of unit  $k$  at period  $t$  is defined by

$$SLPROD_{VA}^k(t, b) \equiv \frac{RVA^k(t, b)}{L^{kt}}. \quad (2.32)$$

It is not unimportant to notice that its dimension is money-of-period- $b$  per unit of labour.

### 2.4.3 Linking Levels and Indices

We now turn to the relation between levels and indices. One expects that taking the ratio of two levels would deliver an index, but let us have a look. Dividing unit  $k$ 's total factor or labour productivity level at period 1 by the same at period 0 delivers, using the various definitions and relations (2.17), (2.12) and (2.16),

$$\frac{TFPROD_{VA}^k(1, b)}{TFPROD_{VA}^k(0, b)} = \frac{Q_{VA}^k(1, b)/Q_{VA}^k(0, b)}{Q_{KL}^k(1, b)/Q_{KL}^k(0, b)}, \quad (2.33)$$

$$\frac{LPROD_{VA}^k(1, b)}{LPROD_{VA}^k(0, b)} = \frac{Q_{VA}^k(1, b)/Q_{VA}^k(0, b)}{Q_L^k(1, b)/Q_L^k(0, b)}, \quad (2.34)$$

$$\frac{SLPROD_{VA}^k(1, b)}{SLPROD_{VA}^k(0, b)} = \frac{Q_{VA}^k(1, b)/Q_{VA}^k(0, b)}{L^{k1}/L^{k0}}, \quad (2.35)$$

respectively. Surely, if  $Q_{VA}^k(t, t')$ ,  $Q_{KL}^k(t, t')$  and  $Q_L^k(t, t')$  are well-behaving functions then the right-hand sides of expressions (2.33), (2.34) and (2.35) have the form of an output quantity index divided by an input quantity index, both for period 1 relative to period 0. When  $b = 0, 1$  one easily checks that (2.33) reduces to  $ITFP_{VA}^k(1, 0)$ , that (2.34) reduces to  $ILPROD_{VA}^k(1, 0)$ , and that (2.35) reduces to  $ISLPROD_{VA}^k(1, 0)$ . But, when  $b \neq 0, 1$ , then

$$TFPROD_{VA}^k(1, b)/TFPROD_{VA}^k(0, b) = ITFP_{VA}^k(1, 0)$$

if and only if the quantity indices  $Q_{VA}^k(t, t')$  and  $Q_{KL}^k(t, t')$  are transitive (that is, satisfy the Circularity Test). Similarly,

$$LPROD_{VA}^k(1, b)/LPROD_{VA}^k(0, b) = ILPROD_{VA}^k(1, 0)$$

if and only if the quantity indices  $Q_{VA}^k(t, t')$  and  $Q_L^k(t, t')$  are transitive, and

$$SLPROD_{VA}^k(1, b)/SLPROD_{VA}^k(0, b) = ISLPROD_{VA}^k(1, 0)$$

if and only if the quantity index  $Q_{VA}^k(t, t')$  is transitive.

Unfortunately, transitive quantity indices are in practice seldom used. Moreover, they would lead to price indices which fail some basic axioms.

#### 2.4.4 When Not All the Data Are Accessible

The word ‘ideally’ was deliberately inserted in front of definitions (2.23), (2.25), (2.26) and (2.27). This word reflects the assumption that all the detailed price and quantity data, necessary to compile production-unit-specific price and quantity index numbers, are accessible. In practice, especially in the case of microdata, though the data are *available* at the enterprises—because revenue and cost are sums of quantities produced or used at certain unit prices—they are usually not *accessible* for researchers, due to the excessive cost of obtaining such data, their confidentiality, the response burden experienced by enterprises, or other reasons. In such cases researchers have to fall back at indices which are estimated for a higher aggregation level. This in turn means that real values are contaminated by differential price developments between the production units considered and the higher level aggregate.

In the extended version of this paper a number of situations are reviewed. In sectoral studies it appears that the way value added is deflated influences the distributions of the ensuing productivity levels; and the same holds at the input side.

A pervasive feature of microdata studies is the use of higher-level instead of production-unit specific deflators. There is some literature on the effect of using industry-level deflators instead of enterprise-level deflators on the estimation of production functions and the analysis of productivity change. See the early study of Abbott (1991) and, more recently, Mairesse and Jaumandreu (2005) and Foster et al. (2008). Of course, for such studies one needs enterprise-level price data, which severely limits the possibilities. In the literature, productivity based on revenue or value added deflated by an industry-level price index is sometimes called ‘revenue productivity’, to distinguish it from our concept that is then called ‘(physical) output productivity’.<sup>6</sup>

In a recent contribution Smeets and Warzynski (2013) found that physical productivity exhibited more dispersion than revenue productivity. A similar feature was unveiled by Eslava et al. (2013). In the last study it was also found that the correlation coefficient of the two measures was low. On the failure of revenue productivity measures to identify within-plant efficiency gains from exporting, see Marin and Voigtländer (2013). From the cross-sectional perspective this issue was studied by van Biesebroeck (2009).

## 2.5 Decompositions: Arithmetic Approach

Let us now assume that productivity levels, real output divided by real input, are somehow available.<sup>7</sup> We denote the productivity level of unit  $k$  at period  $t$  by  $PROD^{kt}$ . Each production unit comes with some measure of relative size (importance) in the form of a weight  $\theta^{kt}$ . These weights add up to 1 for each period, that is

$$\sum_{k \in \mathcal{K}^0} \theta^{k0} = \sum_{k \in \mathcal{K}^1} \theta^{k1} = 1. \quad (2.36)$$

---

<sup>6</sup>The distinction between revenue productivity and physical productivity is a central issue in the microdata study of Hsieh and Klenow (2009), where Indian, Chinese, and U.S. manufacturing plants/firms were compared over the period 1977–2005. However, they didn’t have access to plant/firm-level deflators. Using some theoretical reasoning, real value added was estimated as  $RVA^k(t, b) = (VA^k)^{3/2}$ , so that the ratio of physical productivity, calculated as  $RVA^k(t, b)/X_{KL}^k(t, b)$ , and revenue productivity, calculated as  $VA^k/X_{KL}^k(t, b)$ , becomes equal to  $(VA^k)^{1/2}$  ( $k \in \mathcal{K}^1$ ). It comes as no surprise then that physical productivity exhibits more dispersion than revenue productivity.

<sup>7</sup>This section updates Balk (2003, Sect. 6).

We concentrate here on the productivity levels as introduced in the previous section; that is,  $PROD^{kt}$  has the form of real value added divided by real primary input or real labour input. Then, ideally, the relative size measure  $\theta^{kt}$  must be consistent with either of those measures. Though rather vague, this assumption is for the time being sufficient; we will return to this issue in Sect. 2.9.

The aggregate (or mean) productivity level at period  $t$  is quite naturally defined as the weighted arithmetic average of the unit-specific productivity levels, that is  $PROD^t \equiv \sum_{k \in \mathcal{K}^t} \theta^{kt} PROD^{kt}$ , where the summation is taken over all production units existing at period  $t$ . The weighted geometric average, which is a natural alternative, as well as the weighted harmonic average, will be discussed in the next section.

Aggregate productivity change between periods 0 and 1 is then given by

$$PROD^1 - PROD^0 = \sum_{k \in \mathcal{K}^1} \theta^{k1} PROD^{k1} - \sum_{k \in \mathcal{K}^0} \theta^{k0} PROD^{k0}. \quad (2.37)$$

Given the distinction between continuing, exiting, and entering production units, as defined by expressions (2.10) and (2.11), expression (2.37) can be decomposed as

$$\begin{aligned} PROD^1 - PROD^0 = & \sum_{k \in \mathcal{N}^1} \theta^{k1} PROD^{k1} \\ & + \sum_{k \in \mathcal{C}^{01}} \theta^{k1} PROD^{k1} - \sum_{k \in \mathcal{C}^{01}} \theta^{k0} PROD^{k0} \\ & - \sum_{k \in \mathcal{X}^0} \theta^{k0} PROD^{k0}. \end{aligned} \quad (2.38)$$

The first term at the right-hand side of the equality sign shows the contribution of entering units, the second and third term together show the contribution of continuing units, whereas the last term shows the contribution of exiting units. The contribution of continuing units,  $\sum_{k \in \mathcal{C}^{01}} \theta^{k1} PROD^{k1} - \sum_{k \in \mathcal{C}^{01}} \theta^{k0} PROD^{k0}$ , is the joint outcome of intra-unit productivity change,  $PROD^{k1} - PROD^{k0}$ , and inter-unit relative size change,  $\theta^{k1} - \theta^{k0}$ , for all  $k \in \mathcal{C}^{01}$ . The problem of decomposing this joint outcome into the contributions of the two factors happens to be structurally similar to the index number (or indicator) problem. Whereas in index number theory we talk about prices, quantities, and commodities, we are here talking about sizes, productivity levels, and (continuing) production units.

It can thus be expected that in reviewing the various decomposition methods familiar names from index number theory, such as Laspeyres, Paasche, and Bennet, will turn up (see Balk 2008 for the nomenclature).

### 2.5.1 The First Three Methods

The first method decomposes the contribution of the continuing units into a Laspeyres-type contribution of intra-unit productivity change and a Paasche-type contribution of relative size change:

$$\begin{aligned}
 PROD^1 - PROD^0 = & \\
 & \sum_{k \in \mathcal{N}^1} \theta^{k1} PROD^{k1} \\
 & + \sum_{k \in \mathcal{C}^{01}} \theta^{k0} (PROD^{k1} - PROD^{k0}) + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) PROD^{k1} \\
 & - \sum_{k \in \mathcal{X}^0} \theta^{k0} PROD^{k0}. \tag{2.39}
 \end{aligned}$$

The second term at the right-hand side of the equality sign relates to intra-unit productivity change and uses base period weights. It is therefore, using the language of index number theory, called a Laspeyres-type measure. The third term relates to relative size change and is weighted by comparison period productivity levels. It is therefore called a Paasche-type measure. This decomposition was used in the early microdata study of Baily, Hulten and Campbell (BHC) (1992).

One feature is important to notice. Disregard for a moment entering and exiting production units. Then aggregate productivity change is entirely due to continuing units, and is the sum of two terms. Suppose that all the units experience productivity increase, that is,  $PROD^{k1} > PROD^{k0}$  for all  $k \in \mathcal{C}^{01}$ . Then aggregate productivity change is not necessarily positive, because the relative-size-change term  $\sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) PROD^{k1}$  can exert a negative influence. This ‘paradox’ was extensively discussed by Fox (2012) and we will return to this issue in a later section.

Since base period and comparison period weights add up to 1, we can insert an arbitrary scalar  $a$ , and obtain

$$\begin{aligned}
 PROD^1 - PROD^0 = & \\
 & \sum_{k \in \mathcal{N}^1} \theta^{k1} (PROD^{k1} - a) \\
 & + \sum_{k \in \mathcal{C}^{01}} \theta^{k0} (PROD^{k1} - PROD^{k0}) + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) (PROD^{k1} - a) \\
 & - \sum_{k \in \mathcal{X}^0} \theta^{k0} (PROD^{k0} - a). \tag{2.40}
 \end{aligned}$$

At this point it is useful to introduce some additional notation. Let  $PROD^{\mathcal{X}^0} \equiv \sum_{k \in \mathcal{X}^0} \theta^{k0} PROD^{k0} / \sum_{k \in \mathcal{X}^0} \theta^{k0}$  be the mean productivity level of the exiting units, and let  $PROD^{\mathcal{N}^1} \equiv \sum_{k \in \mathcal{N}^1} \theta^{k1} PROD^{k1} / \sum_{k \in \mathcal{N}^1} \theta^{k1}$  be the mean productivity level of the entering units. Then expression (2.40) can be written as



$$\begin{aligned}
PROD^1 - PROD^0 = & \left( \sum_{k \in \mathcal{N}^1} \theta^{k1} \right) (PROD^{\mathcal{N}^1} - a) \\
& + \sum_{k \in \mathcal{C}^{01}} \theta^{k0} (PROD^{k1} - PROD^{k0}) + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) (PROD^{k1} - a) \\
& - \left( \sum_{k \in \mathcal{X}^0} \theta^{k0} \right) (PROD^{\mathcal{X}^0} - a). \tag{2.41}
\end{aligned}$$

Thus, entering units contribute positively to aggregate productivity change when their mean productivity level exceeds  $a$ , and exiting units contribute positively when their mean productivity level falls short of  $a$ . The net effect of entrance and exit is given by the sum of the first and the fourth right-hand side term,  $(\sum_{k \in \mathcal{N}^1} \theta^{k1})(PROD^{\mathcal{N}^1} - a) - (\sum_{k \in \mathcal{X}^0} \theta^{k0})(PROD^{\mathcal{X}^0} - a)$ . It is interesting to notice that this effect not only depends on relative importancies and mean productivities, but also on the value chosen for the arbitrary scalar  $a$ . However, as we will see, there are a number of reasonable options here.

The second method uses a Paasche-type measure for intra-unit productivity change and a Laspeyres-type measure for relative size change. This leads to

$$\begin{aligned}
PROD^1 - PROD^0 = & \left( \sum_{k \in \mathcal{N}^1} \theta^{k1} \right) (PROD^{\mathcal{N}^1} - a) \\
& + \sum_{k \in \mathcal{C}^{01}} \theta^{k1} (PROD^{k1} - PROD^{k0}) + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) (PROD^{k0} - a) \\
& - \left( \sum_{k \in \mathcal{X}^0} \theta^{k0} \right) (PROD^{\mathcal{X}^0} - a). \tag{2.42}
\end{aligned}$$

I am not aware of any application of this decomposition.

It is possible to avoid the choice between the Laspeyres-Paasche-type and the Paasche-Laspeyres-type decomposition. The third method uses for the contribution of both intra-unit productivity change and relative size change Laspeyres-type measures. However, this simplicity is counterbalanced by the necessity to introduce a covariance-type term:

$$\begin{aligned}
PROD^1 - PROD^0 = & \left( \sum_{k \in \mathcal{N}^1} \theta^{k1} \right) (PROD^{\mathcal{N}^1} - a) \\
& + \sum_{k \in \mathcal{C}^{01}} \theta^{k0} (PROD^{k1} - PROD^{k0}) + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) (PROD^{k0} - a)
\end{aligned}$$

$$\begin{aligned}
& + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0})(PROD^{k1} - PROD^{k0}) \\
& - \left( \sum_{k \in \mathcal{K}^0} \theta^{k0} \right) (PROD^{\mathcal{K}^0} - a).
\end{aligned} \tag{2.43}$$

In view of the overall Laspeyres-type perspective, a natural choice for the arbitrary scalar now seems to be  $a = PROD^0$ , the base period aggregate productivity level. This leads to the decomposition originally proposed by Haltiwanger (1997) and preferred by Foster, Haltiwanger and Krizan (FHK) (2001) (there called method 1). This method has been employed *inter alia* by Foster et al. (2006), Foster et al. (2008), and Collard-Wexler and de Loecker (2013).<sup>8</sup> The FHK method will also be used in OECD's MultiProd project (OECD 2014).

Baldwin and Gu (2006) suggested to set  $a = PROD^{\mathcal{K}^0}$ , the base period mean productivity level of the exiting units. It is clear that then the final right-hand side term in expression (2.43) vanishes, and that the net effect of entrance and exit becomes equal to  $(\sum_{k \in \mathcal{N}^1} \theta^{k1})(PROD^{\mathcal{N}^1} - PROD^{\mathcal{K}^0})$ . It is as if the entering units have replaced the exiting units, and that the mean productivity surplus is all that matters.

Choosing  $a = 0$  brings us back to the BHC decomposition. Nishida et al. (2014) provided interesting comparisons of the BHC and FHK decompositions on Chilean, Colombian and Slovenian micro-level data.

### 2.5.2 Interlude: The TRAD, CSLS, and GEA Decompositions

Let us pause for a while at this expression and consider the case where there is neither exit nor entry; that is  $\mathcal{K}^0 = \mathcal{K}^1 = \mathcal{C}^{01}$ . Then expression (2.43) reduces to

$$\begin{aligned}
PROD^1 - PROD^0 &= \\
& \sum_{k \in \mathcal{C}^{01}} \theta^{k0} (PROD^{k1} - PROD^{k0}) \\
& + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0})(PROD^{k0} - a) \\
& + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0})(PROD^{k1} - PROD^{k0}).
\end{aligned} \tag{2.44}$$

<sup>8</sup>Altomonte and Nicolini (2012) applied the FHK method to aggregate price-cost margin change. For any individual production unit the price-cost margin was defined as nominal cash flow (= value added minus labour cost) divided by nominal revenue,  $CF^{kt}/R^{kt}$ . These margins were weighted by market shares  $R^{kt} / \sum_{k \in \mathcal{K}^t} R^{kt}$  ( $k \in \mathcal{K}^t$ ).

In order to transform to (forward looking) percentage changes (aka growth rates) both sides of this expression are divided by  $PROD^0$ , which delivers

$$\begin{aligned}
 \frac{PROD^1 - PROD^0}{PROD^0} = & \\
 & \sum_{k \in C^{01}} \theta^{k0} \frac{PROD^{k0}}{PROD^0} \left( \frac{PROD^{k1} - PROD^{k0}}{PROD^{k0}} \right) \\
 & + \sum_{k \in C^{01}} (\theta^{k1} - \theta^{k0}) \frac{PROD^{k0} - a}{PROD^0} \\
 & + \sum_{k \in C^{01}} \frac{PROD^{k0}}{PROD^0} (\theta^{k1} - \theta^{k0}) \left( \frac{PROD^{k1} - PROD^{k0}}{PROD^{k0}} \right). \quad (2.45)
 \end{aligned}$$

Now consider *simple* labour productivity, that is, real value added per unit of labour; thus  $PROD^{kt} = SLPROD_{VA}^k(t, b) \equiv RVA^k(t, b)/L^{kt}$  ( $k \in C^{01}$ ). Let the relative size of a production unit be given by its labour share; that is,  $\theta^{kt} \equiv \frac{L^{kt}}{\sum_{k \in C^{01}} L^{kt}}$  ( $k \in C^{01}$ ). It is straightforward to check that then the weights occurring in the first right-hand side term expression (2.45),  $\theta^{k0} \frac{PROD^{k0}}{PROD^0}$ , reduce to real-value-added shares,  $\frac{RVA^k(0, b)}{\sum_{k \in C^{01}} RVA^k(0, b)}$  ( $k \in C^{01}$ ), so that

$$\begin{aligned}
 \frac{PROD^1 - PROD^0}{PROD^0} = & \\
 & \sum_{k \in C^{01}} \frac{RVA^k(0, b)}{\sum_{k \in C^{01}} RVA^k(0, b)} \left( \frac{PROD^{k1} - PROD^{k0}}{PROD^{k0}} \right) \\
 & + \sum_{k \in C^{01}} (\theta^{k1} - \theta^{k0}) \frac{PROD^{k0} - a}{PROD^0} \\
 & + \sum_{k \in C^{01}} \frac{PROD^{k0}}{PROD^0} (\theta^{k1} - \theta^{k0}) \left( \frac{PROD^{k1} - PROD^{k0}}{PROD^{k0}} \right). \quad (2.46)
 \end{aligned}$$

In view of the fact that  $\sum_{k \in C^{01}} (\theta^{k1} - \theta^{k0}) = 0$ , expression (2.46) can also be written as

$$\begin{aligned}
 \frac{PROD^1 - PROD^0}{PROD^0} = & \\
 & \sum_{k \in C^{01}} \frac{RVA^k(0, b)}{\sum_{k \in C^{01}} RVA^k(0, b)} \left( \frac{PROD^{k1} - PROD^{k0}}{PROD^{k0}} \right)
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) \frac{PROD^{k0} - a}{PROD^0} \\
& + \sum_{k \in \mathcal{C}^{01}} \frac{PROD^{k0}}{PROD^0} (\theta^{k1} - \theta^{k0}) \left( \frac{PROD^{k1} - PROD^{k0} - a'}{PROD^{k0}} \right), \quad (2.47)
\end{aligned}$$

for another arbitrary scalar  $a'$ . Now, choosing  $a = 0$  and  $a' = 0$  yields the TRAD(itional) way of decomposing aggregate labour productivity change into contributions of the various industries, according to three main sources: a within-sector effect, a reallocation level effect, and a reallocation growth effect respectively (see for various other names and their provenances De Avillez 2012). Choosing  $a = PROD^0$  and  $a' = PROD^1 - PROD^0$  yields the CSLS decomposition (which has been developed at the Centre for the Study of Living Standards).

Finally, let the relative size of a production unit be given by its combined labour and relative price share; that is,  $\theta^{kt} \equiv \frac{L^{kt}}{\sum_{k \in \mathcal{C}^{01}} L^{kt}} \frac{P_{VA}^k(t, b)}{P_{VA}^k(t, b)}$  ( $k \in \mathcal{C}^{01}$ ), where  $P_{VA}^k(t, b)$  is some non- $k$ -specific deflator. Notice that these weights do not add up to 1. It is straightforward to check that in this case the weights occurring in the first right-hand side term expression (2.45),  $\theta^{k0} \frac{PROD^{k0}}{PROD^0}$ , reduce to nominal-value-added shares,  $\frac{VA^{k0}}{\sum_{k \in \mathcal{C}^{01}} VA^{k0}}$  ( $k \in \mathcal{C}^{01}$ ), so that

$$\begin{aligned}
& \frac{PROD^1 - PROD^0}{PROD^0} = \\
& \sum_{k \in \mathcal{C}^{01}} \frac{VA^{k0}}{\sum_{k \in \mathcal{C}^{01}} VA^{k0}} \left( \frac{PROD^{k1} - PROD^{k0}}{PROD^{k0}} \right) \\
& + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) \frac{PROD^{k0} - a}{PROD^0} \\
& + \sum_{k \in \mathcal{C}^{01}} \frac{PROD^{k0}}{PROD^0} (\theta^{k1} - \theta^{k0}) \left( \frac{PROD^{k1} - PROD^{k0} - a'}{PROD^{k0}} \right), \quad (2.48)
\end{aligned}$$

where  $a$  and  $a'$  are arbitrary scalars. For  $a = 0$  and  $a' = 0$  this appears to be the Generalized Exactly Additive Decomposition (GEAD), going back to Tang and Wang (2004) and explored by Dumagan (2013).

De Avillez (2012) provided an interesting empirical comparison of TRAD, CSLS, and GEAD. He found that “despite some similarities, all three decomposition formulas paint very different pictures of which sectors drove labour productivity growth in the Canadian business sector during the 2000–2010 period.” The difference between TRAD and CSLS not unexpectedly hinges on the role played by the scalars  $a$  and  $a'$ . Varying  $a$  and/or  $a'$  implies varying magnitudes of the two reallocation effects, not at the aggregate level—because the sums are invariant—but at the level of individual production units (in his case, industries).

The difference between TRAD and CSLS on the one hand and GEAD on the other evidently hinges on the absence or presence of relative price levels in the sectoral measures of importance.<sup>9</sup> De Avillez found it “impossible to say which set of estimates provides a more accurate picture of economic reality because the GEAD formula is, ultimately, measuring something very different from the TRAD and CSLS formulas.” I concur insofar this conclusion only means that the answer cannot be found within the bottom-up perspective. The top-down perspective is required to obtain a decision.

### 2.5.3 The Fourth and Fifth Method

Let us now return to expression (2.43). Instead of the Laspeyres perspective, one might as well use the Paasche perspective. The covariance-type term accordingly appears with a negative sign. Thus, the fourth decomposition is

$$\begin{aligned}
 PROD^1 - PROD^0 = & \left( \sum_{k \in \mathcal{N}^1} \theta^{k1} \right) (PROD^{\mathcal{N}^1} - a) \\
 & + \sum_{k \in \mathcal{C}^{01}} \theta^{k1} (PROD^{k1} - PROD^{k0}) + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) (PROD^{k1} - a) \\
 & - \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) (PROD^{k1} - PROD^{k0}) \\
 & - \left( \sum_{k \in \mathcal{X}^0} \theta^{k0} \right) (PROD^{\mathcal{X}^0} - a). \tag{2.49}
 \end{aligned}$$

The natural choice for  $a$  would now be  $PROD^1$ , the comparison period aggregate productivity level. Choosing  $a = PROD^{\mathcal{N}^1}$  would lead to disappearance of the entry effect. The net effect of entrance and exit then becomes equal to  $(\sum_{k \in \mathcal{X}^0} \theta^{k0}) (PROD^{\mathcal{N}^1} - PROD^{\mathcal{X}^0})$ . It is left to the reader to explore the analogs to expressions (2.47) and (2.48) by using backward looking percentage changes. I am not aware of any empirical application of this decomposition.

The fifth method avoids the Laspeyres-Paasche dichotomy altogether, by using the symmetric Bennet-type method. This amounts to taking the arithmetic average of the first and the second method. The covariance-type term then disappears. Thus,

$$PROD^1 - PROD^0 =$$

---

<sup>9</sup>See also Reinsdorf (2015). He considered a convex combination of CSLS with price reference periods  $b = 0$  and  $b = 1$ .

$$\begin{aligned}
& \left( \sum_{k \in \mathcal{N}^1} \theta^{k1} \right) (PROD^{\mathcal{N}^1} - a) \\
& + \sum_{k \in \mathcal{C}^{01}} \frac{\theta^{k0} + \theta^{k1}}{2} (PROD^{k1} - PROD^{k0}) \\
& + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) \left( \frac{PROD^{k0} + PROD^{k1}}{2} - a \right) \\
& - \left( \sum_{k \in \mathcal{X}^0} \theta^{k0} \right) (PROD^{\mathcal{X}^0} - a). \tag{2.50}
\end{aligned}$$

With respect to the scalar  $a$  there are several options available in the literature. A rather natural choice is  $a = (PROD^0 + PROD^1)/2$ , the overall two-period mean aggregate productivity level. Then, entering units contribute positively to aggregate productivity change if their mean productivity level is above this overall mean. Exiting units contribute positively if their mean productivity level is below the overall mean. Continuing units can contribute positively in two ways: if their productivity level increases, or if the units with productivity levels above (below) the overall mean increase (decrease) in relative size. This decomposition basically corresponds to the one used in the early microdata study of Griliches and Regev (GR) (1995). Because of its symmetry it is widely preferred. Moreover, Foster et al. (2001) argue that the GR method (there called method 2) is presumably less sensitive to (random) measurement errors than the asymmetric FHK method. The GR method was employed by Baily et al. (2001) and Foster et al. (2008).

But other choices are also plausible. Baldwin and Gu (BG) (2006) suggested to set  $a = PROD^{\mathcal{X}^0}$ , the base period mean productivity level of the exiting production units. Then, as we have seen before, the last term of expression (2.50) disappears. Put otherwise, entering units are seen as displacing exiting units, contributing positively to aggregate productivity change insofar their mean productivity level exceeds that of the exiting units.

Baldwin and Gu (2008) considered two alternatives, to be applied to different types of industries. The first is to set  $a$  equal to the base period mean productivity level of the continuing units that are contracting; that is, the units  $k \in \mathcal{C}^{01}$  for which  $\theta^{k1} < \theta^{k0}$ . The second is to set  $a$  equal to the base period mean productivity level of the continuing units that are expanding.

Balk and Hoogenboom-Spijker (2003) compared the five methods, defined by expressions (2.41), (2.42), (2.43), (2.49), and (2.50) respectively, on micro-level data of the Netherlands manufacturing industry over the period 1984–1999. Though there appeared to be appreciable differences between the various decompositions, the pervasive fact was the preponderance of the productivity change of the continuing units (or, the ‘within’ term).

### 2.5.4 Another Five Methods

A common feature of the five decomposition methods discussed hitherto is that the productivity levels of exiting and entering production units are compared to a single overall benchmark level  $a$ , for which a number of options is available. It seems more natural to compare the productivity levels of exiting units to the mean level of the continuing units at the base period—which is the period of exit, and to compare the productivity levels of entering units to the mean level of the continuing units at the comparison period—which is the period of entrance.

Thus, let the aggregate productivity level of the continuing production units at period  $t$  be defined as  $PROD^{C^{01}t} \equiv \sum_{k \in C^{01}} \theta^{kt} PROD^{kt} / \sum_{k \in C^{01}} \theta^{kt}$  ( $t = 0, 1$ ). Since the weights  $\theta^{kt}$  add up to 1 for both periods—see expression (2.36)—expression (2.37) can be decomposed as

$$\begin{aligned} PROD^1 - PROD^0 = & \left( \sum_{k \in N^1} \theta^{k1} \right) (PROD^{N^1} - PROD^{C^{01}1}) \\ & + PROD^{C^{01}1} - PROD^{C^{01}0} \\ & - \left( \sum_{k \in X^0} \theta^{k0} \right) (PROD^{X^0} - PROD^{C^{01}0}). \end{aligned} \quad (2.51)$$

This expression tells us that entering units contribute positively to aggregate productivity change if their mean productivity level is above that of the continuing units at the entrance period. Similarly, exiting units contribute positively if their mean productivity level is below that of the continuing units at the period of exit.

Let the relative size of continuing units be defined by  $\tilde{\theta}^{kt} \equiv \theta^{kt} / \sum_{k \in C^{01}} \theta^{kt}$  ( $k \in C^{01}; t = 0, 1$ ). The contribution of the continuing units to aggregate productivity change can then be written as

$$PROD^{C^{01}1} - PROD^{C^{01}0} = \sum_{k \in C^{01}} \tilde{\theta}^{k1} PROD^{k1} - \sum_{k \in C^{01}} \tilde{\theta}^{k0} PROD^{k0}, \quad (2.52)$$

which has the same structure as the second and third term of expression (2.38), the difference being that the weights now add up to 1; that is,  $\sum_{k \in C^{01}} \tilde{\theta}^{kt} = 1$  ( $t = 0, 1$ ). Thus the five methods discussed earlier can simply be repeated on the right-hand side of expression (2.51). The first four, asymmetric, methods are left to the reader. The symmetric Bennet decomposition delivers the following result,

$$\begin{aligned} PROD^1 - PROD^0 = & \left( \sum_{k \in N^1} \theta^{k1} \right) (PROD^{N^1} - PROD^{C^{01}1}) \\ & + \sum_{k \in C^{01}} \frac{\tilde{\theta}^{k0} + \tilde{\theta}^{k1}}{2} (PROD^{k1} - PROD^{k0}) \end{aligned}$$

$$\begin{aligned}
& + \sum_{k \in \mathcal{C}^{01}} (\tilde{\theta}^{k1} - \tilde{\theta}^{k0}) \left( \frac{PROD^{k0} + PROD^{k1}}{2} - a \right) \\
& - \left( \sum_{k \in \mathcal{K}^0} \theta^{k0} \right) (PROD^{\mathcal{K}^0} - PROD^{\mathcal{C}^{01}0}). \tag{2.53}
\end{aligned}$$

The first right-hand side term of this expression refers to the entering production units. As we see, its magnitude is determined by the period 1 share of the entrants and the productivity gap with the continuing units. The last right-hand side term refers to the exiting production units. The magnitude of this term depends on the share of the exiting units and the productivity gap with the continuing units. The second and third term refer to the continuing production units. They may contribute positively in two ways: if their productivity levels on average increase, or if the units with mean productivity levels above (or below) the scalar  $a$  increase (or decrease) in relative size.

Notice that the third term is the only place where an arbitrary scalar  $a$  can be inserted, since the relative weights of the continuing production units add up to 1 in both periods. Though the term itself is invariant, the unit-specific components  $(\tilde{\theta}^{k1} - \tilde{\theta}^{k0})((PROD^{k0} + PROD^{k1})/2 - a)$  are not.

This decomposition was developed by Diewert and Fox (DF) (2010), the discussion paper version of which was published in 2005. Though in a different context—the development of shares of labour types in plant employment—the DF decomposition in the field of productivity measurement can be detected in Vainiomäki (1999). Currently there are hardly any empirical applications known.<sup>10</sup>

If there are no exiting or entering units, that is,  $\mathcal{K}^0 = \mathcal{K}^1 = \mathcal{C}^{01}$ , then the DF method (2.53) as well as the GR method (2.50) reduce to the simple Bennet-type decomposition.

---

<sup>10</sup>Though Kirwan et al. (2012) contend to use the DF decomposition, it appears that their analysis is simply based on expression (2.37). The part relating to continuing units is replaced by a weighted sum of production function based unit-specific productivity changes plus residuals. Böckerman and Maliranta (2007) used the DF decomposition for the analysis of value-added based simple labour productivity and total factor productivity. Kauhanen and Maliranta (2012) applied a two-stage DF decomposition to mean wage change.



### 2.5.5 Provisional Evaluation

The overview provided in the foregoing subsections hopefully demonstrates a number of things, the first and most important of which is that there is no unique decomposition of aggregate productivity change as defined by expression (2.37).<sup>11</sup>

Second, one should be careful with reifying the different components, in particular the covariance-type term, since this term can be considered as a mere artefact arising from the specific (Laspeyres- or Paasche-) perspective chosen.

Third, the undetermined character of the scalar  $a$  lends additional arbitrariness to the first set of five decompositions. At the aggregate level it is easily seen that letting  $a$  tend to 0 will lead to a larger contribution of the entering units, the exiting units, and the size change of continuing units, at the expense of intra-unit productivity change. The advantage of the second set of five decompositions, among which the symmetric DF method, is that the distribution of these four parts is kept unchanged. The remaining arbitrariness in expression (2.53) is in the size-change term and materializes only at the level of individual continuing production units.

Fourth, what counts as ‘entrant’ or ‘exiting unit’ depends not only on the length of the time span between the periods 0 and 1, but also on the length of the periods itself and the observation thresholds employed in sampling.

All in all it can be expected that the outcome of any decomposition exercise depends to some extent on the particular method favoured by the researcher. This is not a problem as long as he or she realizes this and let the favoured results be accompanied by some alternatives.

Finally, as demonstrated in the previous section, the productivity levels  $PROD^{kt}$  depend on the price reference period of the deflators used. In particular this holds for the simple labour productivity levels  $RVA^k(t, b)/L^{kt}$ . This dependence obviously extends to aggregate productivity change  $PROD^1 - PROD^0$ . To mitigate its effect, one considers instead the forward-looking growth rate of aggregate productivity  $(PROD^1 - PROD^0)/PROD^0$  and its decomposition, obtained by dividing each term by  $PROD^0$ . It would of course be equally justified to consider the backward-looking growth rate  $(PROD^1 - PROD^0)/PROD^1$ . A symmetric growth rate is obtained when the difference  $PROD^1 - PROD^0$  is divided by a mean of  $PROD^0$  and  $PROD^1$ . When the logarithmic mean<sup>12</sup> is used, one obtains

$$\frac{PROD^1 - PROD^0}{L(PROD^0, PROD^1)} = \ln(PROD^1 / PROD^0), \quad (2.54)$$

<sup>11</sup>This non-uniqueness should not come as a surprise and finds its parallel in index number theory (see Balk 2008) and in so-called structural decomposition analysis (widely used in input-output analysis; see Dietzenbacher and Los 1998).

<sup>12</sup>The logarithmic mean, for any two strictly positive real numbers  $a$  and  $b$ , is defined by  $L(a, b) \equiv (a - b) / \ln(a/b)$  if  $a \neq b$  and  $L(a, a) \equiv a$ . See Balk (2008, pp. 134–136) for a discussion of its properties, one of these being that  $(ab)^{1/2} \leq L(a, b) \leq (a + b)/2$ .

which can be interpreted as a percentage change.<sup>13</sup> However, its decomposition still contains differences such as  $PROD^{k1} - PROD^{k0}$ , which of course can be transformed into logarithmic differences but at the expense of getting pretty complicated weights.

Thus this calls for going geometric right from the start; which is the topic of the next section.

## 2.6 Decompositions: Geometric and Harmonic Approach

In the geometric approach the aggregate productivity level is defined as a weighted *geometric* average of the unit-specific productivity levels, that is  $PROD^t \equiv \prod_{k \in K^t} (PROD^{kt})^{\theta^{kt}}$ . This is equivalent to defining  $\ln PROD^t \equiv \sum_{k \in K^t} \theta^{kt} \ln PROD^{kt}$ , which implies that, by replacing  $PROD$  by  $\ln PROD$ , the entire story of the previous section can be repeated.

The advantage of decomposing  $\ln PROD^1 - \ln PROD^0$  over decomposing  $PROD^1 - PROD^0$  is that a logarithmic change can be interpreted immediately as a percentage change. The disadvantage is that, as an aggregate level measure, a geometric mean  $\prod_{k \in K^t} (PROD^{kt})^{\theta^{kt}}$  is less easy to understand than an arithmetic mean  $\sum_{k \in K^t} \theta^{kt} PROD^{kt}$ . We let the top-down approach here advise which mean should be preferred; see the extended version of this paper.

The Geometric DF decomposition was applied by Hyytinen and Maliranta (2013). They extended the decomposition to deal with age groups of firms. There are of course also geometric versions of the GR, FHK, and BG decompositions. Baldwin and Gu (2011) compared these on Canadian retail trade and manufacturing industry microdata over the period 1984–1998. As in the comparative study of Balk and Hoogenboom-Spijker (2003), they found that in manufacturing the ‘within’ term was dominant. However, in retail trade the net effect of entry and exit appeared more important.

In the harmonic approach the aggregate productivity level is defined as a weighted *harmonic* average of the unit-specific productivity levels, that is  $PROD^t \equiv (\sum_{k \in K^t} \theta^{kt} (PROD^{kt})^{-1})^{-1}$ . Though the literature does hardly pay any attention to this option, in the extended version of this paper it will be shown that there are situations in which this type of average rather naturally emerges. An example is provided by Böckerman and Maliranta (2012). Though these authors were primarily concerned with the evolution of the aggregate real labour share through time, it turns out that their analysis is equivalent to an Harmonic DF decomposition on aggregate labour productivity, defined as weighted harmonic mean of  $LPROD_{VA}^k(t, t-1)$  with weights defined as real value added shares at period  $t$ .

---

<sup>13</sup>Since  $\ln(a/a') = \ln(1 + (a - a')/a') \approx (a - a')/a'$  when  $(a - a')/a'$  is small.

## 2.7 Monotonicity

As already preluded to, all the definitions of aggregate productivity change, whether arithmetic, geometric or harmonic, suffer from what Fox (2012) called the “monotonicity problem” or “paradox”.

Again, disregard for a moment entering and exiting production units. Then aggregate productivity change is entirely due to continuing units, and is the combination (sum, product, or harmonic sum, respectively) of two terms. Suppose that all the units experience productivity increase, that is,  $PROD^{k1} > PROD^{k0}$  for all  $k \in C^{01}$ . Then the ‘within’ term in the DF decomposition (2.53) and in the Geometric DF decomposition is positive, and in the Harmonic DF decomposition negative. However, aggregate productivity change is not necessarily positive, because the relative-size-change terms, can exert a counterbalancing influence.

Fox (2012) noticed that the term  $\sum_{k \in C^{01}} (\theta^{k0} + \theta^{k1})(PROD^{k1} - PROD^{k0})/2$  as such has the desired monotonicity property, and proposed to extend this measure to the set  $C^{01} \cup \mathcal{X}^0 \cup \mathcal{N}^1 = \mathcal{K}^0 \cup \mathcal{N}^1 = \mathcal{X}^0 \cup \mathcal{K}^1$ . Aggregate productivity change is then defined as

$$\Delta PROD_{Fox}(1, 0) \equiv \sum_{k \in C^{01} \cup \mathcal{X}^0 \cup \mathcal{N}^1} \frac{\theta^{k0} + \theta^{k1}}{2} (PROD^{k1} - PROD^{k0}). \quad (2.55)$$

Now, for all exiting production units,  $k \in \mathcal{X}^0$ , it is evidently the case that in the later period 1 those units have size zero; that is,  $\theta^{k1} = 0$ . It is then rather natural to set their virtual productivity level also equal to zero; that is,  $PROD^{k1} = 0$ . Likewise, entering units,  $k \in \mathcal{N}^1$ , have size zero in the earlier period 0; that is,  $\theta^{k0} = 0$ . Their virtual productivity level at that period is also set equal to zero; that is,  $PROD^{k0} = 0$ . Then expression (2.55) can be decomposed as

$$\begin{aligned} \Delta PROD_{Fox}(1, 0) &= \\ &= (1/2) \sum_{k \in \mathcal{N}^1} \theta^{k1} PROD^{k1} \\ &+ \sum_{k \in C^{01}} \frac{\theta^{k0} + \theta^{k1}}{2} (PROD^{k1} - PROD^{k0}) \\ &- (1/2) \sum_{k \in \mathcal{X}^0} \theta^{k0} PROD^{k0}. \end{aligned} \quad (2.56)$$

Unfortunately, there is no geometric or harmonic analog to expressions (2.55) and (2.56), because the logarithm or reciprocal of a zero productivity level is infinity. By using logarithmic means, one obtains

$$\begin{aligned}
\Delta PROD_{Fox}(1, 0) = & \\
& (1/2) \sum_{k \in \mathcal{N}^1} \theta^{k1} PROD^{k1} \\
& + \sum_{k \in \mathcal{C}^{01}} \frac{\theta^{k0} + \theta^{k1}}{2} L(PROD^{k0}, PROD^{k1}) \ln (PROD^{k1} / PROD^{k0}) \\
& - (1/2) \sum_{k \in \mathcal{X}^0} \theta^{k0} PROD^{k0}, \tag{2.57}
\end{aligned}$$

which, however, does not provide any advantage vis à vis expression (2.56).

It is interesting to compare Fox's proposal to the symmetric decomposition (2.50) with  $a = 0$ . It turns out that

$$\begin{aligned}
PROD^1 - PROD^0 = & \\
& \Delta PROD_{Fox}(1, 0) \\
& + (1/2) \sum_{k \in \mathcal{N}^1} \theta^{k1} PROD^{k1} \\
& + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) \frac{PROD^{k0} + PROD^{k1}}{2} \\
& - (1/2) \sum_{k \in \mathcal{X}^0} \theta^{k0} PROD^{k0}. \tag{2.58}
\end{aligned}$$

It is remarkable that of the entire contribution of entering and exiting production units to  $PROD^1 - PROD^0$ , half is considered as productivity change and half as non-productivity change. It is difficult to envisage a solid justification for this.

## 2.8 The Olley-Pakes Decomposition

Though aggregate, or weighted mean, productivity levels are interesting, researchers are also interested in the *distribution* of the unit-specific levels  $PROD^{kt}$  ( $k \in \mathcal{K}^t$ ), and the change of such distributions over time, a good example being Bartelsman and Dhrymes (1998). Given the relative size measures  $\theta^{kt}$ —which are adding up to 1—a natural question is whether high or low productivity of a unit goes together with high or low size. Are big firms more productive than small firms? Or are the most productive firms to be found among the smallest? Questions multiply when the time dimension is taken into account. Does the ranking of a particular production unit in the productivity distribution sustain through time? Are firms ranked somewhere in a particular period likely to rank higher or lower in the next period? Is there a relation

with the age, however determined, of the production units? Do the productivity distributions, and the behaviour of the production units, differ over the industries?

When it comes to size a natural measure to consider is the covariance of weights and productivity levels. Let  $\#(\mathcal{K}^t)$  be the number of units in  $\mathcal{K}^t$ , let  $\overline{PROD}^t \equiv \sum_{k \in \mathcal{K}^t} PROD^{kt} / \#(\mathcal{K}^t)$  be the unweighted mean of the productivity levels, and let  $\bar{\theta}^t \equiv \sum_{k \in \mathcal{K}^t} \theta^{kt} / \#(\mathcal{K}^t) = 1 / \#(\mathcal{K}^t)$  be the unweighted mean of the weights. One then easily checks that

$$\sum_{k \in \mathcal{K}^t} (\theta^{kt} - \bar{\theta}^t)(PROD^{kt} - \overline{PROD}^t) = PROD^t - \overline{PROD}^t. \quad (2.59)$$

This is a particular instance of a general relation derived by Bortkiewicz in 1923/1924. Bortkiewicz showed that the difference between two differently weighted means has the form of a covariance. Interesting applications can be found in index number theory (see Balk 2008).

Olley and Pakes (OP) (1996, 1290) rearranged this relation to the form

$$PROD^t = \overline{PROD}^t + \sum_{k \in \mathcal{K}^t} (\theta^{kt} - \bar{\theta}^t)(PROD^{kt} - \overline{PROD}^t) \quad (2.60)$$

and provided an interpretation which has been repeated, in various forms, by many researchers.<sup>14</sup> The interpretation usually goes like this: There is some event (say, a certain technological innovation or some other shock) that gives rise to a productivity level  $\overline{PROD}^t$ ; but this productivity level is transformed into an aggregate level  $PROD^t$  by means of a mechanism called *reallocation*, the extent of which is measured by the covariance term in expression (2.60). So it seems that the aggregate productivity level  $PROD^t$  is ‘caused’ by two factors, a productivity shock and a reallocation.<sup>15</sup>

I propose to call this the Olley-Pakes *fallacy*, because there are not at all two factors. Expression (2.59) is a mathematical identity: reallocation, defined as a covariance, is identically equal to the difference of two means, a weighted and an unweighted one. All that expression (2.60) does is featuring the unweighted mean rather than the weighted mean as the baseline variable.

I don’t dispute the usefulness of studying time-series or cross-sections of covariances such as we see at the left-hand side of expression (2.59). Notice that by replacing  $PROD$  by  $\ln PROD$  or  $1/PROD$  one obtains a geometric or harmonic variant respectively. Additional insight can be obtained when one replaces produc-

<sup>14</sup>It is straightforward to generalize the OP decomposition to the case where the ensemble  $\mathcal{K}^t$  consists of a number of disjunct groups. The right-hand side of expression (2.60) then becomes the sum of a between-groups covariance and for each group an unweighted mean productivity level and a within-group covariance. Collard-Wexler and de Loecker (2013) considered a case of two groups.

<sup>15</sup>In Foster et al. (2001)’s article the OP decomposition, expression (2.60), was called method 3.

tivity levels  $PROD^{kt}$  by productivity changes, measured as differences  $PROD^{k1} - PROD^{k0}$  or percentage changes  $\ln(PROD^{k1}/PROD^{k0})$ . As a descriptive device this is wonderful, especially for comparing ensembles (industries, economies)—see for instance Lin and Huang (2012) where such covariances are regressed on several background variables. In the cross-country study of Bartelsman et al. (2013) within-industry covariances between size and productivity play a key role.<sup>16</sup>

The OP decomposition, expression (2.60), can of course be used to decompose aggregate productivity change  $PROD^1 - PROD^0$  into two terms, the first being  $\overline{PROD}^1 - \overline{PROD}^0$ , and the second being the difference of two covariance terms. But then we are unable to distinguish between the contributions of exiting, continuing, and entering production units. Thus, it is advisable to restrict the OP decomposition to the continuing units, and substitute into expression (2.51). Doing this results in the following expression,

$$\begin{aligned}
 PROD^1 - PROD^0 = & \left( \sum_{k \in \mathcal{N}^1} \theta^{k1} \right) (PROD^{\mathcal{N}^1} - PROD^{C^{01}1}) \\
 & + \overline{PROD}^{C^{01}1} - \overline{PROD}^{C^{01}0} \\
 & + \sum_{k \in C^{01}} (\tilde{\theta}^{k1} - 1/\#(C^{01})) (PROD^{k1} - \overline{PROD}^{C^{01}1}) \\
 & - \sum_{k \in C^{01}} (\tilde{\theta}^{k0} - 1/\#(C^{01})) (PROD^{k0} - \overline{PROD}^{C^{01}0}) \\
 & - \left( \sum_{k \in \mathcal{X}^0} \theta^{k0} \right) (PROD^{\mathcal{X}^0} - PROD^{C^{01}0}), \tag{2.61}
 \end{aligned}$$

where  $PROD^{C^{01}t}$  is the weighted mean productivity level and  $\overline{PROD}^{C^{01}t} \equiv \sum_{k \in C^{01}} PROD^{kt} / \#(C^{01})$  is the unweighted mean productivity level of the continuing units at period  $t$  ( $t = 0, 1$ );  $\#(C^{01})$  is the number of those units.

This then is the decomposition proposed by Melitz and Polanec (2015). Their paper contains an interesting empirical comparison of the GR method (2.50), the FHK method (2.43), and the extended OP method (2.61). Hansell and Nguyen (2012) compared the BG method (2.50), the DF method (2.53), and the extended OP method (2.61). Again, their overall conclusion on Australian data concerning the 2002–2010 period was that the “dominant source of labour productivity growth in manufacturing and professional services is from within firms.”

<sup>16</sup>See also the special issue on “Misallocation and Productivity” of the *Review of Economic Dynamics* 16(1)(2013). There appears to be no unequivocal definition of ‘misallocation’. In OECD (2014) at least three different concepts can be detected.

Wolf (2011, pp. 21–25), see also Bartelsman and Wolf (2014), used the OP decomposition to enhance the GR decomposition. By substituting expression (2.60) into expression (2.50), with  $a = (PROD^0 + PROD^1)/2$ , one obtains

$$\begin{aligned}
 PROD^1 - PROD^0 = & \left( \sum_{k \in \mathcal{N}^1} \theta^{k1} \right) \left( PROD^{\mathcal{N}^1} - \frac{PROD^0 + PROD^1}{2} \right) \\
 & + \sum_{k \in \mathcal{C}^{01}} \frac{\theta^{k0} + \theta^{k1}}{2} (PROD^{k1} - PROD^{k0}) \\
 & + \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) \left( \frac{PROD^{k0} + PROD^{k1}}{2} - \frac{\overline{PROD}^0 + \overline{PROD}^1}{2} \right) \\
 & - \sum_{k \in \mathcal{C}^{01}} (\theta^{k1} - \theta^{k0}) \left( \sum_{k \in \mathcal{K}^0} (\theta^{k0} - \bar{\theta}^0) (PROD^{k0} - \overline{PROD}^0) \right. \\
 & \left. + \sum_{k \in \mathcal{K}^1} (\theta^{k1} - \bar{\theta}^1) (PROD^{k1} - \overline{PROD}^1) \right) / 2 \\
 & - \left( \sum_{k \in \mathcal{X}^0} \theta^{k0} \right) \left( PROD^{\mathcal{X}^0} - \frac{PROD^0 + PROD^1}{2} \right). \tag{2.62}
 \end{aligned}$$

As one sees, the original GR ‘between’ term, the third right-hand side term in expression (2.50), is split into two parts. The first part, which is the third right-hand side term in the last expression, is relatively easy to understand: it is still a covariance between size changes and mean productivity levels. The second part, which is the fourth right-hand side term in the last expression, is far more complex. This part can be rewritten as  $(\sum_{k \in \mathcal{N}^1} \theta^{k1} - \sum_{k \in \mathcal{X}^0} \theta^{k0})$  times a mean covariance (of size and productivity level). It is unclear how this could be interpreted.

## 2.9 The Choice of Weights

The question which weights  $\theta^{kt}$  are appropriate when a choice has been made as to the productivity levels  $PROD^{kt}$  ( $k \in \mathcal{K}^t$ ) has received some attention in the literature. Given that somehow  $PROD^{kt}$  is output divided by input, should  $\theta^{kt}$  be output- or input-based? And how is this related to the type of mean—arithmetic, geometric, or

harmonic? The literature does not provide us with definitive answers.<sup>17</sup> Indeed, as long as one stays in the bottom-up framework it is unlikely that a convincing answer can be obtained. We need the complementary top-down view.

A bit formally, the problem can be posed as follows. Generalizing the three definitions used in Sects. 2.5 and 2.6, aggregate productivity is a weighted ‘mean’ of individual productivities

$$PROD^t \equiv M(\theta^{kt}, PROD^{kt}; k \in \mathcal{K}^t), \quad (2.63)$$

where the ‘mean’  $M(\cdot)$  can be arithmetic, geometric, or harmonic; the weights  $\theta^{kt}$  may or may not add up to 1; and  $PROD^{kt}$  can be value added based  $TFPROD$ ,  $LPROD$  or  $SLPROD$ , as defined in Sect. 2.4.1, or gross-output based  $TFPROD$  or  $SLPROD$ , as defined in the extended paper. The task then is: find the set of weights such that

$$PROD^t = PROD^{\mathcal{K}^t t}, \quad (2.64)$$

that is, such that aggregate productivity can be interpreted as productivity of the aggregate.

It is clear that there are a number of options here, but the discussion of these can be found in the extended version of this paper.

## 2.10 Conclusion

The main lessons can be summarized as follows:

1. Generically, productivity is defined as output over input. Yet most, if not all, empirical studies are not about productivity as such, because there is contamination by price effects at the input and/or at the output side of the production units considered. In many sectoral studies the available deflators are more or less deficient; for instance, value added is single-deflated instead of double-deflated. In almost all microdata studies there are simply no firm- or plant-specific deflators available and higher-level substitutes must be used instead. All this may or may not matter at the aggregate (industry or economy)

---

<sup>17</sup>As Karagiannis (2013) showed, the issue is not unimportant. He considered the OP decomposition (2.60) on Greek cotton farm data. Output and input shares were used to weight total factor productivity and labour productivity levels. The covariances turned out to be significantly different. An earlier example was provided by van Beveren (2012), using firm-level data from the Belgian food and beverage industry. de Loecker and Konings (2006) noted that there is no clear consensus on the appropriate weights (shares) that should be used. In their work they used employment based shares  $L^{kt} / \sum_k L^{kt}$  to weight value-added based total factor productivity indices  $Q_{VA}^k(t, b) / Q_{KL}^k(t, b)$ .



level, but it does matter when it comes to judging the contribution of specific (sets of) production units to aggregate productivity (change).

2. Economists appear to have a preference for working with levels; e.g. with concepts such as real value added. It is good to realize, as pointed out in Sect. 2.4, that a level actually is a long-term index. And this implies that there is always some, essentially arbitrary, normalization involved. For instance, there is a time period for which real value added equals nominal value added; or, there is a period for which total factor productivity equals 1.
3. Essentially the bottom-up approach consists in aggregating micro-level productivities with help of some set of size-related weights and then decomposing aggregate productivity change into contributions of (specific sets of) continuing, entering, and exiting units. We have seen that there is a large number of such decompositions available. Because of its symmetry and its natural benchmarks for exiting and entering production units we prefer the Diewert-Fox decomposition.
4. Beware of the covariance, so-called “reallocation”, terms; e.g. in expressions (2.43), (2.49), or (2.60). They are statistical artefacts and there is not necessarily some underlying economic process involved.
5. In the bottom-up approach not every combination of micro-level productivities, weights, and aggregator function leads to a nice interpretation of aggregate productivity as productivity of the aggregate. The complementary top-down approach should be our guide here. The connection between the two approaches is discussed in the extended version of this paper.

## References

- Abbott TA (1991) Producer price dispersion, real output, and the analysis of production. *J Prod Anal* 2:179–195
- Altomonte C, Nicolini M (2012) Economic integration and the dynamics of firms’ competitive behavior. *Struct Chang Econ Dyn* 23:383–402
- Baily MN, Hulten C, Campbell D (1992) Productivity dynamics in manufacturing plants. *Brook Pap Econ Act Microecon* 2:187–249
- Baily MN, Bartelsman EJ, Haltiwanger J (2001) Labor productivity: structural change and cyclical dynamics. *Rev Econ Stat* 83:420–433
- Baldwin JR, Gu W (2006) Plant turnover and productivity growth in Canadian manufacturing. *Ind Corp Chang* 15:417–465
- Baldwin JR, Gu W (2008) Firm turnover and productivity growth in the Canadian retail trade sector. *Economic Analysis (EA) research paper series*, vol 53. Statistics Canada, Ottawa
- Baldwin JR, Gu W (2011) Firm dynamics and productivity growth: a comparison of the retail trade and manufacturing sectors. *Ind Corp Chang* 20:367–395
- Balk BM (2003) The residual: on monitoring and benchmarking firms, industries, and economies with respect to productivity. *J Prod Anal* 20:5–47
- Balk BM (2008) *Price and quantity index numbers: models for measuring aggregate change and difference*. Cambridge University Press, New York
- Balk BM (2010) An assumption-free framework for measuring productivity change. *Rev Income Wealth* 56(Special Issue 1):S224–S256

- Balk BM (2014) Dissecting aggregate output and labour productivity change. *J Prod Anal* 42:35–43
- Balk BM (2015) Measuring and relating aggregate and subaggregate total factor productivity change without neoclassical assumptions. *Statistica Neerlandica* 69:21–28
- Balk BM, Hoogenboom-Spijker E (2003) The measurement and decomposition of productivity change: exercises on the Netherlands' manufacturing industry. Discussion Paper 03001. Statistics Netherlands, Den Haag
- Balk BM, Reich UP (2008) Additivity of national accounts reconsidered. *J Econ Soc Meas* 33:165–178
- Bartelsman EJ, Dhrymes PJ (1998) Productivity dynamics: US manufacturing plants, 1972–1986. *J Prod Anal* 9:5–34
- Bartelsman EJ, Doms M (2000) Understanding productivity: lessons from longitudinal microdata. *J Econ Lit* XXXVIII:569–594
- Bartelsman EJ, Wolf Z (2014) Forecasting aggregate productivity using information from firm-level data. *Rev Econ Stat* 96:745–755
- Bartelsman EJ, Haltiwanger J, Scarpetta S (2013) Cross-country differences in productivity: the role of allocation and selection. *Am Econ Rev* 103:305–334
- van Beveren I (2012) Total factor productivity estimation: a practical review. *J Econ Surv* 26:98–128
- van Biesebroeck J (2009) Disaggregate productivity comparisons: sectoral convergence in OECD countries. *J Prod Anal* 32:63–79
- Böckerman P, Maliranta M (2007) The micro-level dynamics of regional productivity growth: the source of divergence in Finland. *Reg Sci Urban Econ* 37: 165–182
- Böckerman P, Maliranta M (2012) Globalization, creative destruction, and labour share change: evidence on the determinants and mechanisms from longitudinal plant-level data. *Oxford Econ Pap* 64: 259–280
- Collard-Wexler A, de Loecker J (2013) Reallocation and technology: evidence from the US steel industry. Working Paper 18739. National Bureau of Economic Research, Cambridge
- De Avillez R (2012) Sectoral contributions to labour productivity growth in Canada: does the choice of decomposition formula matter? *Int Prod Monit* 24:97–117
- Dietzenbacher E, Los B (1998) Structural decomposition techniques: sense and sensitivity. *Econ Syst Res* 10:307–323
- Diewert WE, Fox KJ (2010) On measuring the contribution of entering and exiting firms to aggregate productivity growth. In: Diewert WE, Balk BM, Fixler D, Fox KJ, Nakamura AO (eds) *Price and productivity measurement. Index number theory*, vol 6. Trafford Press, Victoria. [www.vancouvervolumes.com](http://www.vancouvervolumes.com). [www.indexmeasures.com](http://www.indexmeasures.com). Revised version of Discussion Paper 05–02. Department of Economics, University of British Columbia, Vancouver (2005)
- Dumagan JC (2013) A generalized exactly additive decomposition of aggregate labor productivity growth. *Rev Income Wealth* 59:157–168
- Dumagan JC, Balk BM (2015) Dissecting aggregate output and labour productivity change: a postscript on the role of relative prices. *J Prod Anal* (online)
- Durand R (2004) Uniqueness of the numeraire and consistent valuation in accounting for real values. *J Econ Soc Meas* 29: 411–426
- Eslava M, Haltiwanger J, Kugler A, Kugler M (2013) Trade and market selection: evidence from manufacturing plants in Colombia. *Rev Econ Dyn* 16:135–158
- Foster L, Haltiwanger J, Krizan CJ (2001) Aggregate productivity growth: lessons from microeconomic evidence. In: Hulten CR, Dean ER, Harper MJ (eds) *New developments in productivity analysis. Studies in income and wealth*, vol 63. The University of Chicago Press, Chicago/London
- Foster L, Haltiwanger J, Krizan CJ (2006) Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s. *Rev Econ Stat* 88:748–758
- Foster L, Haltiwanger J, Syverson C (2008) Reallocation, firm turnover, and efficiency: selection on productivity or profitability? *Am Econ Rev* 98:394–425

- Fox KJ (2012) Problems with (dis)aggregating productivity, and another productivity paradox. *J Prod Anal* 37:249–259
- Griliches Z, Regev H (1995) Firm productivity in Israeli industry, 1979–1988. *J Econ* 65:175–203
- Haltiwanger J (1997) Measuring and analyzing aggregate fluctuations: the importance of building from microeconomic evidence. *Fed Reserve Bank St. Louis Econ Rev* 79(3):55–77
- Hansell D, Nguyen T (2012) Productivity, entry and exit in Australian manufacturing and professional services. Paper presented at economic measurement group workshop, University of New South Wales, Sydney, 21–23 November 2012
- Hsieh CT, Klenow PJ (2009) Misallocation and manufacturing TFP in China and India. *Q J Econ* 124:1403–1448
- Hyytinen A, Maliranta M (2013) Firm lifecycles and evolution of industry productivity. *Res Policy* 42:1080–1098
- Karagiannis G (2013) Reallocation and productivity dynamics: empirical evidence on the role of aggregation weights. Mimeo. Department of Economics, University of Macedonia, Thessaloniki
- Kauhanen A, Maliranta M (2012) Micro-components of aggregate wage dynamics. ETLA Working papers No. 1. The Research Institute of the Finnish Economy.
- Kirwan BE, Uchida S, White TK (2012) Aggregate and farm-level productivity growth in tobacco: before and after the quota buyout. *Am J Agric Econ* 94:838–853
- Lin YC, Huang TH (2012) Creative destruction over the business cycle: a stochastic frontier analysis. *J Prod Anal* 38:285–302
- de Loecker J, Konings J (2006) Job reallocation and productivity growth in a post-socialist economy: evidence from Slovenian manufacturing. *Eur J Polit Econ* 22:388–408
- Mairesse J, Jaumandreu J (2005) Panel-data estimates of the production function and the revenue function: what difference does it make? *Scand J Econ* 107:651–672
- Marin AG, Voigtländer N (2013) Exporting and plant-level efficiency gains: it's in the measure. Working Paper 19033. National Bureau of Economic Research, Cambridge
- Melitz MJ, Polanec S (2015) Dynamic Olley-Pakes productivity decomposition with entry and exit. *Rand J Econ* 46: 362–375
- Nishida M, Petrin A, Polanec S (2014) Exploring reallocation's apparent weak contribution to growth. *J Prod Anal* 42:187–210
- OECD (2014) The micro drivers of aggregate productivity. Report DSTI/EAS/IND/WPIA(2014)4. Organisation for Economic Co-operation and Development, Paris
- Olley S, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64:1263–1297
- Reinsdorf M (2015) Measuring industry contributions to labour productivity change: a new formula in a chained Fisher index framework. *Int Prod Monit* 28: 3–26
- Smeets V, Warzynski F (2013) Estimating productivity with multi-product firms, pricing heterogeneity and the role of international trade. *J Int Econ* 90:237–244
- Syverson C (2011) What determines productivity? *J Econ Lit* 49:326–365
- Tang J, Wang W (2004) Sources of aggregate labour productivity growth in Canada and the United States. *Can J Econ* 37:421–444
- Vainiomäki J (1999) Technology and skill upgrading: results from linked worker-plant data for Finnish manufacturing. In: Haltiwanger JC, Lane JJ, Spletzer JR, Theeuwes JJM, Troske KR (ed) *The creation and analysis of employer-employee matched data. Contributions to Economic Analysis* 241. Emerald Bingley
- Wolf Z (2011) Aggregate productivity growth under the microscope. Research Series, vol 518. Tinbergen Institute, VU University Amsterdam

Productivity and Efficiency Analysis

Greene, W.H.; Khalaf, L.; Sickles, R.; Veall, M.; Voia, M.-C.  
(Eds.)

2016, XII, 331 p. 50 illus., 31 illus. in color., Hardcover

ISBN: 978-3-319-23227-0